Isolation and Complete Sequence of a Functional Human Glyceraldehyde-3-phosphate Dehydrogenase Gene*

(Received for publication, March 14, 1988)

Louis Ercolani[‡], Brian Florence, Maria Denaro, and Maria Alexander

From the Department of Medicine and Howard Hughes Medical Institute Laboratories, Harvard Medical School and the Medical Services and Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts 02114

Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.2.12) (GAPDH) mRNA levels, protein, and enzymatic activity increase in 3T3-F442A adipocytes after exposure to physiological concentrations of insulin (Alexander, M., Curtis, G., Avruch, J., and Goodman, H. (1985) J. Biol. Chem. 260, 11978-11985). In order to understand the mechanism of this regulation, we have isolated and sequenced 5.4 kilobase pairs of a 12kilobase pair human genomic clone encoding a functional GAPDH gene. The gene consists of 9 exons and 8 introns with eukaryotic signals necessary for the transcription and translation of GAPDH mRNA. The exon sequence confirms previously published cDNA sequences for human GAPDH in muscle, liver, and erythrocytes. The organization of the human and the unique chicken GAPDH genes is strikingly similar. Although chicken exons VIII-XI have been fused into human exon 8, introns which separate exons encoding the NAD binding, catalytic, and helical domains of the GAPDH protein have been retained. Stable transfection of rodent cells with the intact human GAPDH gene resulted in the expression of a correctly initiated human GAPDH mRNA and an enzymatically active human GAPDH polypeptide. Thus, the gene contains a functional promoter and intact coding sequences. Although many processed GAPDH pseudogenes and GAPDH-like sequences are present in the human genome, Southern blot analysis of human genomic DNA using a probe derived from the 3'-untranslated region of the GAPDH gene detected only two genes, a 10-copy processed pseudogene and a single copy of the isolated gene. In contrast, a probe derived from an intron segment of the isolated gene detected only a single copy of the GAPDH gene. Collectively, these findings strongly suggest that the human genome encodes a single functional GAPDH gene.

The glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.2.12) (GAPDH)1 catalyzes the conversion of glyceraldehyde 3-phosphate to 1,3-diphosphoglycerate (1-3). GAPDH is a tetramer composed of identical subunits of

37,000 molecular weight (2, 3). The activity of the enzyme is regulated by several glycolytic metabolites and by the association of the enzyme with cell membranes (4, 5). Our interest in GAPDH stems from recent observations that insulin increases GAPDH gene transcription in 3T3-F442A adipocytes² which results in a proportionate increase in GAPDH mRNA levels and enzymatic activity in these cells (6). This effect and other effects of insulin on diverse metabolic processes is initiated through the interaction of insulin with a ligandspecific cell surface receptor with intrinsic tyrosine kinase activity (7). The events subsequent to this interaction that result in alterations in gene expression have yet to be elucidated. As current concepts suggest that alterations in gene transcription may result from a direct interaction of transacting factors with cis-acting sequences in hormone-regulated genes, utilizing the sequences responsible for insulin-mediated regulation of this gene should allow us to isolate and identify trans-acting factors that mediate the effect of insulin on GAPDH gene transcription. Thus, we sought to isolate a functional GAPDH gene.

Although somatic hybrid studies have localized a functional GAPDH gene to chromosome 12 and no isoenzymatic forms of the gene with glycolytic activity have been described (8), recent reports have demonstrated that GADPH belongs to a large multiple gene family. This gene family contains 150 or more GAPDH-like sequences, many of which are processed pseudogenes (9, 10). Thus far, attempts to clone the human GAPDH gene have resulted in the isolation of two processed pseudogenes for GAPDH localized to the X chromosome (11, 12). In this paper, the primary structure of a 5.37-kilobase segment of genomic DNA containing the functional human GAPDH gene is presented.

EXPERIMENTAL PROCEDURES AND RESULTS³

DISCUSSION

This report describes the isolation and characterization of a human genomic DNA segment which contained a functional GAPDH gene. The isolated gene satisfied our screening criteria insofar as it is the only DNA segment in the human genome that contained both introns and sequences similar to the extreme 3'-untranslated region of the expressed human GAPDH mRNA (see Fig. 3). As seen in Fig. 8, lanes A and B, Southern blot analysis of human genomic DNA performed with a probe derived from intron B of the isolated gene detected one copy of the GAPDH gene. In contrast, a probe

^{*} The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) J04038.

Supported by National Institutes of Health Grant AM7028.

[‡] Supported by National Institutes of Health Grant Alvi 1020.

The abbreviations used are: GAPDH, glyceraldehyde-3-phosphate dehydrogenase; kb, kilobases; pghGAPDH, 12-kb human generated the statement of the st nomic segment encoding GAPDH subcloned into pUC12; SDS, sodium dodecyl sulfate.

 ² M. Alexander and H. M. Goodman, unpublished results.
 ³ Portions of this paper (including "Experimental Procedures," "Results," Figs. 1-7, and Table I) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

Fig. 8. Southern blot analysis of the human GAPDH gene. The structure of the human GAPDH gene is shown with restriction enzyme sites used for Southern analysis: open boxes correspond to the exon segments in the human GAPDH gene or the human GAPDH cDNA. Closed bars indicate intron segments in the GAPDH gene. DNA isolated from human placenta was digested with restriction enzymes and subjected to Southern blot analysis as described under "Experimental Procedures." S corresponds to the intron fragment used to hybridize human placental DNA digested with PstI in lane A and HindIII-BglI in lane B. S corresponds to the 5'-cDNA GAPDH fragment used to hybridize human placental DNA digested with PstI in lane C and HindIII-BglI in lane D. I corresponds to the 3'cDNA GAPDH fragment used to hybridize placental DNA digested with HindIII-BglI in lane E and PstI in lane F. Z corresponds to the 3'-flanking fragment of the human GAPDH gene used to hybridize placental DNA digested with PstI in lane G and HindIII-BglI in lane H.

derived from the 3'-untranslated region of this gene detected one copy of the GAPDH gene and 10 copies of a processed GAPDH pseudogene (lanes G and H). We believe this pseudogene is the same DNA segment cloned by Hanauer and Mandel (12) because the exon portion of our 3'-flanking probe was sufficient in length (84 bases) and in sequence similarity (96%) to detect this pseudogene. This finding indicated that there are only two genes in the human genome with sequences similar to the 3'-untranslated region of the GAPDH gene expressed in human tissues, the gene we have isolated and a processed pseudogene. For this reason, we are certain that we have cloned the unique GAPDH gene.

Furthermore, the size of the active transcriptional unit was consistent with that predicted by Dani et al. (41) for the human GAPDH gene expressed in HeLa cells. An active promoter was located within 112 base pairs 5' to the start site of transcription of this gene. The mRNA encoded by this human DNA segment was correctly initiated (see Fig. 4). In studies published elsewhere, we have shown that this mRNA directed the synthesis of a full length polypeptide (38). Here we have demonstrated that after stable integration and expression of the gene in H35 hepatoma cells, an enzymatically active human GAPDH polypeptide was produced (see Fig. 5). Thus, we have isolated a functional GAPDH gene.

Nevertheless, we and others (9-12) have detected a very large number of other GAPDH-related sequences in the human genome with probes derived from the GAPDH cDNA

coding region (see Fig. 8). Several groups have speculated about the functional significance of these genes. Acari et al. (36) have suggested that a fetal form of the GAPDH gene exists; however, their report was not confirmed by Fort et al. (34). The origin of the other GAPDH-related sequences remains unclear. It is likely that some may represent segments of other genes, such as kinases and dehydrogenases, which have acquired NAD binding or catalytic domains similar to those in the GAPDH gene through exon shuffling (42, 43). Alternatively, others may represent genes with structures similar to GAPDH, that perform different functions.

By analyzing glycolytic enzyme activity in somatic hybrids, Bruns et al. (8) localized the GAPDH gene to chromosome 12. However, in certain tissues, GAPDH is not a cytosolic protein and appears to mediate functions other than glycolysis. For instance, Tsai et al. (5) demonstrated that GAPDH bound to Band 3, the anion transport protein in erythrocyte cell membranes, is enzymatically inactive. Furthermore, GAPDH may have novel functions which are necessary for the formation of specialized membrane structures in brain (44) and skeletal muscle transverse tubules (45). For instance, Kawamoto and Caswell (46) have suggested that in rabbit skeletal muscle, GAPDH may function as a protein kinase. While glycolytic isoenzymes of GAPDH have not been found (8), the presence of a GAPDH protein which performs diverse functions in certain tissues raised the possibility that other GAPDH genes might exist in the human genome. However,

the recent cloning of p37, an intracellular membrane-associated protein in human erythrocytes which is expressed on the extracellular surface of K562 erythroleukemia cells, unexpectedly revealed this protein to be identical in sequence with the cytosolic form of GAPDH (37). Furthermore, comparison of the human GAPDH gene exon sequence to the combined sequences of GAPDH cDNAs isolated from human adult liver, muscle, and erythrocyte suggested complete identity (see Fig. 3). Thus, it appears the GAPDH gene we have cloned encodes the GAPDH protein that mediates both glycolytic and nonglycolytic functions in all human tissues.

GAPDH is present in both prokaryotes and eukaryotes and is highly conserved. Such conservation makes the GAPDH gene a useful model for the study of evolutionay function of intervening sequences. Based on their studies of the chicken GAPDH gene sequence, Stone et al. (13) have proposed that the ontogeny for this highly conserved protein is the result of duplication events involving ancestral introns as well as exons. The conservation of introns separating domains predicted by these authors to occur in GAPDH genes of higher eukaryotes has been confirmed. The location of three intervening sequences which separate exons encoding four protein domains of the GAPDH protein detected by x-ray crystallography has been retained in the human GAPDH gene (47, 48). Specifically, intron IV in the chicken gene (intron D in the human gene) divides the two mononucleotide-binding domains; intron VI in the chicken (intron F in the human gene) separates the dinucleotide-binding domain (amino acids 1-149) from the catalytic domain (amino acids 149-312); and intron XI in the chicken gene (intron H in the human gene) separates the helical domain (amino acids 313-334).

Three introns in the human gene corresponding to chicken introns VIII-X have been lost in tandem. Random intron deletions have been described in higher eukaryotic genes (49-51). However, the finding of a 10-copy processed GAPDH pseudogene suggested a possible recombinant event between a retroviral cDNA and a GAPDH gene similar in structure to the chicken GAPDH gene. Based on a predicted mutation rate of 0.7% per million years in the nonfunctional GAPDH pseudogene sequence (51), this event would have occurred within the past 5.7 million years. Thus, our hypothesis would predict an intron/exon organization similar to that of the chicken GAPDH gene in primates such as baboons which diverged from pre-human primates before the GAPDH pseudogene was produced, whereas random deletions of these introns could have occurred at any time after the avian radiation of 270 million years ago.

The 5'- and 3'-flanking regions of this gene contain many repetitive elements. Eleven direct repeats were found in the 5'-flanking region ranging from 7 to 19 bases long. Data base searches revealed several sequences that may contribute to the regulation of the human GAPDH gene. Inverted repetitive elements CCGCCC-GGCGGG associated with upstream enhancer elements for the human and mouse metallothionein genes (52-54) were present in position -972 to -950. A core enhancer sequence GTGGAAG described by Weiher et al. (55) was present on the noncoding strand of intron G. "CAAT" box sequences are usually located within 100 bases of the TATAAA sequence (56). A "CAAT" box sequence GGGCCAATCT, present in virtually all globin genes, was located 203 bases upstream of the TATAAA sequence in the human GAPDH gene. However, analysis of constructs with deletions of the 5'-flanking sequences had no more than a 50% decrease in promoter activity when the CAAT sequence was removed.

In studies published elsewhere, we have shown that the

human GAPDH gene is regulated by insulin when transfected and expressed in differentiated F442A adipocytes but not preadipocytes (38). The differentiation of preadipocytes to adipocytes is stimulated by agents which demethylate chromatin (57, 58). It is therefore of some interest that the GAPDH promoter was found in an area (-630 to -1 nucleotides) that is 65% G-C-rich, with 41 CpG dinucleotides and 59 GpC dinucleotides. This area is similar to the well described HpaII tiny fragment islands that are characteristic of mammalian housekeeping genes (59). It is possible that this HpaII tiny fragment island with 41 potential methylation sites may regulate GAPDH gene expression during the differentiation of preadipocytes. The availability of a probe which detects a unique GAPDH gene rather than the numerous pseudogenes present in the mammalian genome will make it possible to investigate whether methylation influences GAPDH gene expression during differentiation.

Having isolated the human GAPDH gene, it will be feasible to delineate the sequences and characterize the factors responsible for the transcriptional activation of this gene by insulin. These studies are currently in progress in our laboratory. We believe the isolation of the human GAPDH gene will afford many opportunities for understanding both gene expression in general and the final effector pathways for insulin action in particular.

Acknowledgments-We are indebted to H. M. Goodman in whose laboratory the isolation of the human GAPDH gene was initiated. We gratefully acknowledge M. J. Siciliano for providing advice on GAPDH isoenzyme assays, C. Shih for providing the human genomic library, Stephanie Quamo for excellent technical assistance, Paul Habermann for sequencing the human GAPDH cDNA, J. R. Grove, R. Selden, B. Seed, and D. Moore for careful reading of the manuscript and helpful comments. L. E. remains indebted to Joseph Avruch for past and continuing support.

REFERENCES

- 1. Dandliker, W. B., and Fox, J. B. (1955) J. Biol. Chem. 214, 275-283
- 2. Fox, J. B., and Dandliker, W. B. (1956) J. Biol. Chem. 218, 53-57
- 3. Harris, I., and Perham, R. N. (1963) Biochem. J. 89, 60P
- 4. Wang, C., and Alapovic, P. (1980) Arch. Biochem. Biophys. 205, 136-145
- 5. Tsai, I.-H., Murthy, S. N. P., and Steck, T. L. (1982) J. Biol. Chem. 257, 1438-1442
- 6. Alexander, M., Curtis, G., Avruch, J., and Goodman, H. (1985) J. Biol. Chem. 260, 11978-11985
- Rosen, O. (1987) Science 237, 1452-1458
- 8. Bruns, G. A. P., Pierce, P., Regina, V. M., and Gerald, P. S. (1978) Cytogenet. Cell Genet. 22, 547-551
- Piechaczyk, M., Blanchard, J. M., Sabouty, S. R.-E., Dani, C., Marty, L., and Jeanteur, P. (1984) Nature 312, 469-471

 10. Tso, J. Y., Sun, X.-H., Kao, T., Reece, K. S., and Wu, R. (1985)
- Nucleic Acids. Res. 13, 2485-2502
- 11. Benham, F. J., Hodgkinson, S., and Davies, K. E. (1984) EMBO J. 3, 2635-2640
- 12. Hanauer, A., and Mandel, J. L. (1984) EMBO J. 3, 2627-2633
- Stone, E. M., Rothblum, K. N., and Schwartz, R. J. (1985) Nature 313, 498-500
- 14. Dugaiczyk, A., Haron, J. A., Stone, E. M., Dennison, O. E., Rothblum, K. N., and Schwartz, R. J. (1983) Biochemistry 22, 1605-1613
- 15. Michelson, A. M., Markham, A. F., and Orkin, S. H. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 472-476
- 16. Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977)
- J. Mol. Biol. 113, 237-251
 17. Shih, C., Burke, C., Chou, M. J., Seldis, J. B., Yang, C. S., Lee, C. S., Isselbacher, K. J., Wands, J. R., and Goodman, H. M. (1987) J. Virol. **61**, 3491–3498
- 18. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517
- 19. Perron-Yanisch, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103-119

- 20. Sinha, N. D., Biernat, J., McManus, J., and Koster, H. (1984) Nucleic Acids Res. 12, 4539-4557
- Chen, E. J., and Seeburg, P. H. (1985) DNA (NY) 4, 165-170
- 22. Seed, B., and Aruffo, A. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 3365-3369
- 23. Devereux, J., Haeberli, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 387-395
- 24. Southern, P. J., and Berg, P. (1982) Mol. Appl. Genet. 1, 327-341
- 25. Gorman, C., Moffat, L., and Howard, B. (1982) Mol. Cell. Biol. 2, 1044-1051
- 26. Lopata, M. A., Cleveland, D. W., and Sollner-Webb, B. (1984) Nucleic Acids Res. 12, 5707-5717
- 27. Selden, R. F., Howie, K. B., Rowe, M. E., Goodman, H. M. and Moore, D. (1986) Mol. Cell. Biol. 6, 3173-3179
- Wright, D. A., Siciliano, M. J., and Baptist, J. N. (1972) Experientia (Basel) 281, 889
- 29. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132.
- 30. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manuel, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 31. Mount, S. M. (1982) Nucleic Acids Res. 10, 459-472
- Dierks, P., Van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J., and Weissman, C. (1983) Cell 32, 695-706
- 33. Cochran, M. D., and Weissmann, C. (1984) EMBO J. 3, 2453-2459
- Fort, P., Marty, L., Piechaczyk, M., Sabrouty, S. E., Dani, C., Jeanteur, P., and Blanchard, J. M. (1985) Nucleic Acids Res. 13. 1431-1442
- Kozak, M. (1987) Nucleic Acids Res. 15, 8125–8147
- 36. Arcari, P., Martinelli, R., and Salvatore, F. (1984) Nucleic Acids Res. 12, 9179-9189
- 37. Allen, R. W., Trach, K. A., and Hoch, J. A. (1987) J. Biol. Chem. **262**, 649-653
- Alexander, M., Lomanto, M., Nasrin, N., and Ramaika, C. (1988)
 Proc. Natl. Acad. Sci. U. S. A. 85, 5092-5096
- 39. Shih, M.-C., Lazar, G., and Goodman, H. M. (1986) Cell 47, 73-
- 40. Stone, E. M., Rothblum, K. N., Alevy, M. C., Kuo, T. M., and Schwartz, R. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 1628-

- 41. Dani, C., Piechaczyk, M., Audiger, Y., El Sabouty, A., Cathala, G., Marty, L., Fort, P., Blanchard, J.-P., and Jeanteur, P. (1984) Eur. J. Biochem. 145, 299-304
- 42. Blake, C. (1985) Int. Rev. Cytol. 93, 149-185
- 43. Duester, G., Jornvall, H., and Hatfield, G. W. (1986) Nucleic Acids Res. 14, 1931-1941
- 44. Huitorel, P., and Pantaloni, D. (1985) Eur. J. Biochem. 150, 265 - 269
- 45. Caswell, A. H., and Corbett, A. M. (1985) J. Biol. Chem. 260, 6892-6898
- 46. Kawamoto, R. M., and Caswell, A. H. (1986) Biochemistry 25, 656-661
- 47. Rossman, M., Liljas, A., Branden, C., and Banaszak, L. (1975) in The Enzymes (Bayer, P. D., ed) vol. 11, pp. 61-102, Academic Press, Orlando, FL
- 48. Harris, J., and Waters, M. (1975) in The Enzymes (Boyer, P. D.,
- ed) Vol. 13, pp. 1-49, Academic Press, Orlando, FL 49. Zakut, R., Shani, M., Givol, D., Neuman, S., Yaffe, D., and Nudel, U. (1982) *Nature* **298**, 857-859
- 50. deCrombrugghe, B., and Pastan, I. (1982) Trends Biochem. Sci. 7, 11
- 51. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner,
- R., and Dogson, J. (1980) Cell 20, 555-556 52. Stuart, G. W., Searle, P. F., Chen, H. Y., Brinster, R. L., and Palmiter, R. D. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 7318-
- 53. Carter, A. D., Felber, B. K., Walling, M., Jubier, M.-F., Schmidt, C. J., and Hamer, D. H. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 7392-7396
- 54. Karin, M., Haslinger, A., Holtgreve, H., Richards, R. I., Krauter, P., Westphal, H. M., and Beato, M. (1984) Nature 308, 513-
- 55. Weiher, H., Konig, M., and Gruss, P. (1983) Science 219, 626-
- 56. Buchler, P., and Trifonov, E. N. (1986) Nucleic Acids Res. 14, 10009-10026
- Sager, R., and Kovac, P. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 480-484
- Chapman, A. B., Knight, D. M., Dieckmann, B. S., and Ringold, G. M. (1984) J. Biol. Chem. 259, 15548-15555
- 59. Bird, A. (1986) Nature 321, 209-213

Supplemental Material to: Isolation and Complete Sequence of a Functional Human Glyceraldehyde:3-Phosphate Dehydrogenase Gene

Louis Ercolani, Brian Florence, Maria Denaro, and Maria Alexander

EXPERIMENTAL PROCEDURES

Cloning the Human GAPDH Gene: A chicken GAPDH cDNA (14) was used to screen a human liver cDNA library (15). A full length human GAPDH cDNA clone was obtained; the insert was excised and purified twice by Seeplaque agarose gel electrophoresis. The fragment was nick-translated by the melthod of Ripby, et al. (16) and used to screen 1 x 10⁶ plaques from an EMBL Sau3A partial library (17). Fifteen positive clones were subjected to restriction endonuclease mapping and Southern analysis (18). A 12 kb insert from a clone with a restriction map consistent with the presence of introns, and a hybridization patient consistent with the presence of the entire human GAPDH coding sequence was subcloned into pUC12 (ppticAPPDH[19]) and used for nucleotide sequencing analysis, and expression studies as described below.

Nucleis Acid Saguencina: Restriction fragments of pghGAPDH were subcloned into either pUC12 (mp18) or M13 (mp18 or mp19) vectors. Oligonucleolide primers 15 to 17 nucleotides in length were synthesized by (i) ps phosphoramidile method on an automated Applied Biosystem 308B DNA synthesized (20). Double stranded pUC12 templates were atkail denatured and ethanol precipitated as described by Chen and Seoburg (21). Singles stranded that stranded charatured templates were heated to 100°C for singles in the presence of the appropriate oligonucleotide primer. The annealing eaction was allowed to proceed for 20 mins at a temperature calculated to be 5°C below the 1m of the heteroduplexes. The heteroduplexes were extended in the presence of [35-3] alpha dATP using the Klenow tragment of DNA polymerase or reverse transcriptase as described by Seed and Aruffo (22). The seguencing reaction mixtures were denatured by healing to 100°C for 2 mins and subjected to electrophoress on prefereded 5-10°C, polyacrylamide gels containing 42% urea. Gels were dired and autoratiographicd with Kodak XAR lim at 23°C for 12-96 hrs. Overlapping sequences were assemblud and analyzed using University of Wisconsin Gonotics Group Software (23).

Cellular Translection Assays: Ten µg of pghGAPDH was cotranslected with 1 µg of pSV2neo (24) into H35 hepatoma cells (a gift of John Koontz, University of Tennessee) by the calcium phosphate precipitation method of Gorman, et al. (25), H35 hepatoma cells containing stably integrated plasmids were selected by resistance to 300 µg /mi Geneticin (G418). Cell extracts were prepared for Western blot analysis, GAPDH enzymatic activity, and primor eviension analysis as described below.

5: flanking sequences of pghGAPDH were subcloned into pOGH, a promotoriess plasmid containing a growth hormone reporter gene (a gift of Richard Selden, Massachusetts General Hospital). These constructs were transiently transfected into L cells using the DEAE dextran method (26) as modified by Selden, et al.(27). Growth hormone secreted into the growth modia of transfected cells was measured 24-48 hrs after transfection using a radioimmunoassay kii (Nichols). The lowest reproducible level of growth hormone detected is 0.1 ng/ml. DNA uplake was normalized by co-transfection with RSV-CAT, a plasmid containing the chloramphenicol acetyl transferase reporter gene driven by the flous Sacroma Virus promoter (26). Results are expressed as % increase ± SEM, in ng/ml growth hormone secreted of constructs vs pOGH, normalized for chloramphenicol acetyl transferase activity.

Primer_Extension_Analysis: The cell monolayer was washed twice with phosphato-bullered saline, and total cytoplasmic RNA was extracted from these monolayers as previously described (6). Identification of the start site of transcription for rat and human GAPDH mRNA was determined by annealing 10 µg of cellular RNA with a [32P]-end labelled oligonocleoide complementary to the coding region of both rat and GAPDH inRNA (5°CCATGTAGTTGAGGTCARTG-3). The RNA-ONA displexes were extended in the presence of cash addition actions on products were extracted with priendly precipitated in the presence of 0.3M sodium acotato and 65% othanol. After resuspension in distilled water the extension products were heat demanued at 100°C for 2 mins and subjected to electrophoresis on preferedud 8% polyacrylamide gels containing 42% urea, dried and subjected to autoradiography at -80°C for 16 hrs. The extension products for rat GAPDH mRNA and human GAPDH mRNA were predicted to be 195 nucleotides and 205 nucleotides in length, respectively.

GAPPH Activitie: The cell monolayer was washed twice with phosphate bullered saline and scraped into four volumes of 10 mM Tris-HCI, pH 7.4, 1 mM EDTA, and 1 mM 2-mercaptoethanol. The cells were homogenized by repeated passage through a 27 gauge needle. The supernatant was prepared by centrifugation at 100,000 x g 10 60 mins at 4°C. The protein contents of the supernatant fractions were determined by Blorad assay. Ten µl of cell extract supernatant mached for protein was subjected to electrophoresis (10 mA/cm for 6 hrs at 0°C) on non-denaturing 1% Soakem agarose or 14% starch horizontal stab gels which contained 8.7 mM Tris-HCI, 2.9 mM sodium citrate, 0.67 mg/mi NAO at pH 7.0. The electrode buffer contained 0.13 M Tris-HCI, 0.043 M sodium citrate at pH 7.0 which was rapidly recirculated to prevent changes in pH. Following electrophorasis, GAPDH tetramers anzymes separated by gel electrophorasis were scanned on at LKB densitometer.

1

₫

Southern Bigutings: Human placental genomic DNA, GAPDH cDNA, and cloned genomic lagments were digested with appropriate restriction endonucleases, separated by electrophoresis on a 1% agarose get and transferred to NEN GeneScreen Plus membranes as described by the manufacturer (Now England Nuclear). Membranes were prehybridized for 24 hrs at 42°C in the presence of 1% SDS. 1 M NaCl. 10% destran sulfate, and 50% desionized formamide. Hybridization was preformed under similar conditions in the presence of denatured probes labelled with [32P] alpha ATP by priming with random hexamers followed by extension of these primines with the Klenow fragment of DNA polymerase (29). After hybridization, membranes were washed for 30 mins, truce in: 2x SSC (1x SSC = 0.15 M NaCl/0.015 M sodium citrate) at 23°C; 2x SSC/1% SDS at 65°C; then 0.1x SSC at 23°C. The membranes were direct and autoradiographed with Kodak XAH film at minus 80°C for 12.96 hrs. The number of GAPDH-related sequences in the human genome was estimated by comparison of the signal strength of genomic DNA fragments to that of multiple dilutions of cloned cDNA fragments.

isolation and Sequence Analysis of Putative Human GAPDH Genomic Clones

A human genomic library was screened with a full length human GAPDH cDNA probe.

A human genomic library was screened with a full length human GAPDH cDNA probe.

Of the 105 recombinant phage screened, 63 positives were detected, 15 of which were analyzed by restriction endonuclesse mapping. One of these clones lacked introns and was clearly a processed pseudogene. Three independent clones had similar maps compatible with the presence of introns. Southern blot analysis of these clones with oligonucleotide probes specific for the 5 and 3 untranslated regions of the human GAPDH cDNA suggested that one of these clones encompassed the entire coding region and one kilobase of 5 flanking sequence in a 12 kb DNA fragment. The insert of this clone was subcloned into the Bamfl site of pDIC12 (pphGAPDH). Limited sequence analysis of pphGAPDH confirmed the presence of introns, a TATAA box, a translation start and slop site, and a polyadonylation signal, which suggested this clone encoded a functional gene. For further nucleotide sequence analysis and expression studies, subclones of the pphGAPDH insert were constructed based on the restriction endonuclease map seen in Fig. 1. Nucleotide sequence was obtained on both strands by sequencing overlapping subclones or sequencing the original pphGAPDH plasmid with oligonucleotide primers. Ambiguous sequences from G-C rich regions were resolved by use of reverse frenerciplase.

As seen in Fig. 2 and Table 1, the Isolated GAPDH gene consisted of 9 exons and 8 introns. The 5 and 3 end of the introns conformed to the sukaryotic gene splice donor/acceptor consensus sequence AACI GTRAGAN ... //6-11/A (31). Typical eukaryotic signals for RNA polymerase II, TATAAA, and polydeonylation, AATAAA were present. The 5 and 3 lialanking regions of this gene were GC inch, 65% and 55% respectively. An additional conserved promoter sequence GGCCAATCT, the "CAAT box" found in virtually all globin genes, was found 203 bases upstream of the buman GAPDH house confirmed by the other four sequences. The

The Isolated Gene Contains a Functional Promoter and Encodes a Functional

The Isolated Gene Contains a Functional Fromoter and Electrons & Functional Prohypoptide.

Restriction tragments of the human GAPDH gene containing the following Stanking septiments: 4871/20, 2684/20, and -112 /120 were subcloned into pOGH, a promoterless plasmid which contains coding sequences from the human growth hormone gene (27). The construct from -288 to +20 included both TATAA and CAAT box sequences, while that from -112 to +20 was devoid of the CAAT box sequence. These constructs were transiently transfected into I. cells. In three experiments, cells

translacted with constructs containing the GAPDH 5-lianking sequences extending to 112, 288, and 487 bases demonstrated a 3178% ± 180, 5038% ± 92, and 5222% ± 1142 increase in growth horizone socretion respectively, as compared to the pOGH plasmid alone which socreted 1.1 ± 0.16 ng/ml of growth horizone. These experiments indicated that the isotated GAPDH gene contained a functional promoter within +112 nucleotides of the putative transcription start site and that deteition of the CAAT box sequence had only a modest effect on gene expression.

The phGAPDH plasmid containing the isotated human GAPDH gene, was stably translacted into rat H35 hepatoma cells. One translacted cell line, 731 was used for the experiments doscribed below. Total RNA was extracted from human placenta, human JEG3 choriocarcinoma cells, rat H35 hepatoma cells, and 731 cells for primer extension product for placenta or JEG3 choriocarcinoma cells was 205 nucleatides (tanes b and c) whereas the extension product rom H35 hepatoma cells was 195 nucleotides (tanes d and e). In contrast, two distinct extension products identical in length to the extension products for human and ral GAPDH mRNA was 195 nucleotides (tanes b and c) whereas the extension product rom H35 hepatoma cells was 195 nucleotides (tanes b and c) products for human and ral GAPDH mRNA was correctly initiated from the plasmid pgliGAPDH. Densitometric analysis of an appropriately exposed autoradiogram revealed 731 cells contained 5-fold more human GAPDH mRNA than rat GAPDH mRNA.

In studies published elsewhere (38), we have demonstrated that I full longth contained activity of this human GAPDH polypepide, cell extractis were subjected to electrophoreoris on non-denaturing starch slab gels, Fig. 5, (lanes A-C) or agarcse gels (tanes D-E) at pl17.0. Under these conditions, GAPDH enzymes from H35 hepatoma colls marginate towards the cathode (lane A). Cells translected with pphGAPDH contained GAPDH enzyme signal active chimenic totramers (lane C). Chimene GAPDH enzymes could be resolv

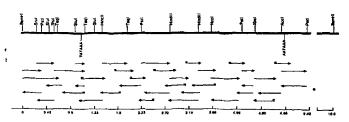
Comparison to the Chicken GAPDH Gene.

Primordial birds and mammals diverged approximately 270 million years ago (39), nevertheless comparison of the human and chicken GAPDH gene sequences revealed striking organizational similarities (40). As seen in Fig. 6, the arrangement of zons differed only in the fusion of chicken exons VIH-XI into human exon 8. Exon lengths between the two genes were virtually identical from Exon 3 to Exon 8 (Tablo 1). The overall length of the two genes was similar differing only by 70 bases in length from the transcriptional start site to the polyadenylation signal. This difference was due to the creation of a larger intron 8 in the human sequence which is not entirely companisated by the intron losses that produced a fused exon 8. As seen in Fig. 7, dot plot comparison of the entire chicken CAPDH gene to the human GAPDH gene revealed striking sequence similarity between exons but only limited sequence similarities between introns and the 5-flanking regions of the two genes. However, the immediate intron-exon borders did show significant similarity which may reflect a contextual conservation of those segments necessary for organization of the splicing machinery in both organisms (Table 1).

Southern Blot Analysis of the Human GAPDH Gene.

Tso, et al. (10)and Piechaczyk, et al. (9) have demonstrated that a targe number of GAPDH-related sequences exist in the human genome. To determine whether the isolated GAPDH DNA segment was unique, human placental DNA was digested with either Hindflif and Bglu, or Pstl and transferred to Gene Serven-Pius membranes for Southern blot analysis. As seen in Fig. 8, a probe derived from intron. B detected a 1.9 kb band in placental DNA digested with Pstl (lane B). The sizes of these detected fragments are identical to those predicted based on the restriction map for this particular GAPDH gene. The signal intensity of in each digest indicated that one copy of this gene was present in the human genome. The same results were obtained with hymphocyte DNA isolated from a different individual (data not shown).

To ascertain whether a GAPDH gene existed in the human genome, Southern blot analysis of human placental DNA digested with Pstl was performed with a probe which contained 60 nucleotides 5: of AATAAA and the 3'-flenking region of the gene (lane G). Only one copy of the predicted 1.1 kb GAPDH gene tragment was detected in Pstl digests of placental DNA in higher molecular weight band corresponding to a copy of a gene with introns dividing Exon 8 was not found. Approximately 10 copies of an apparent 12 kb GAPDH seculated DNA segment was a processed pseudopene. Su this length human GAPDH cDNA was digested with Hindli and fragments 5' or 3' to this site were used to construct the processed pseudopene. In confirm that the 1.2 kb GAPDH-related DNA segment was a processed pseudopene. Can et plot cDNA regements detected the 1.2 kb processed pseudopene. In confirm that the 1.2 kb GAPDH-related DNA dependent with the 3'-untranslated region of the gene (lane G), the pattern of bands detected with the 5' and 3'-GAPDH cDNA probes was markedly more complex (lanes C), C, E, E. These findings indicated that that three distinct types of GAPDH related sequences existed in the human geno



Sequencing strategy for the human GAPDH gene.
 Restriction sites that were utilized to prepare subclones for DNA sequencing are.
 Arrows indicate the direction and extent of individual sequencing runs. Arrows boxes indicate where unique oligonucleotide primers were used to confirm

1136					
	GGATCCCCTGCTGGGAGGGGGCACGGGGACCTGTTCCCACCGTGTGCCCAAGACCTCTTT	(5' PLANK)	a:	CTCTCTGCTCCTCCTGTTCGACAGTCACCCGCATCTTCTTTTGCGTCGCCAGCCGAGCCA	60
-1016	CCCACTTTTTCCCTCTTCTTGACTCACCCTGCCCTAATATCCCCCGGCGCAGCAGTGAA AGGGAGTCCCTGGCTCCTGGCTCQCCTGCACGTCCCAGGGGGGGGGG		C:		60
956	CACGTCCCGCTCTTCGCCCCAGGCTGGATGGAATGAAAGGCACACTGTCTCTCTC		d:		
- 896	GCAGCACAGCCCACAGGTTTCAGGAGTGCCTTTGTGGGAGGCCTCTGGGCCCCCACCAGC		e:		
- 836 - 776	CATCCTGTCCTCCGCCTGGGGCCCCAGCCCGGAGAGAGGCCCGCTGGTGCACACAGGGCCCGG		r:	T A T A A	
- 716	GATTGTCTGCCTAATTATCAGGTCCAGGCTACAGGGCTGCAGGACATCGTGACCTTCCG TGCAGAAACCTCCCCCTCCCCCTCAAGCCGCCTCCGGAGCCTCCTTCCT		a:	Ol Moranes as a second	
- 656	CAGTGCCCAGTGCCCAGTGCCCAGCCCAGGCCTCGGTCCCAGAGATGCCAGGAGCCAGGA		e:	CATCGCTCAGACACCATGGGGAAGGTGAAGGTCGGAGTCAACGGATTTGGTCGTATTGGG	120
- 596	GATGCCGAGGGGAAGTGGGGGCTGGGAAGGAACCACGGGCCCCGGCCCGAGCCCATGGG		f:		
- 536 - 476	CCCCTCCTAGGCCTTTGCCTGAGCAGACCGGTGTCACTACCGCAGAGCCTCGAGGAGAAG TTCCCCAACTTTCCCGCCTCTCAGCCTTTGAAAGAAAGAA		g:	MetGlyLysValLysValGlyValAsnGlyshaGlyArgllaGly	
- 416	TOCASCCOCGACCOGTCCTGGGCTCCGGCTCCAATTCCCCATCTCAGTCGTTCCCAAAGT		-	**************************************	
- 356	CCTCCTGTTTCATCCAAGCGTGTAAGGGTCCCCGTCCTTGACTCCCTAGTGTCCTGCTGC		a :	CGCCTGGTCACCAGGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTTGCCATCAATCA	160
- 296	CCACAGTCCAGTCCTGGGAACCAGCACCGATCACCTCCATCGGGCCAATCXCAGTCCCTT		g:	ArgleuValThrArgAlaAlaFheAsnSerGlyLysValAspTleValAlaIleAsnAsp	140
- 236	CCCCCTACGTCGGGGCCCACACGCTCGGTGCGTGCCCAGTTGAACCAGGCGGCTGCGGAA				
- 176 - 116	AAAAAAAGCGGGGGAGAAAGTAGGGCCGGGCTACTAGCGGTTTTACGGGCGCACGTAGCT CAGGCCTCAAGACCTTGGGCTGGGACTGGCTGAGCCTGGCGGGGGGGG		a:	CCCTTCATTGACCTCAACTACATGGTTTACATGGTCCAATATGATTCCACCCATGGCAAA	240
- 56	CCGCCTGCCGCCGCGCCCCGGTTTCTATAAATTGAGCCCGGAGCCTCCCGCTTCG/		f:	G C A	
	+1		g:	ProPhelleAspleuAsnTyrMetValTyrMetPheGlnTyrAspMerThrHisGlyLys	
		[EXON 1]	a :	TEGOS TO COMPANY AND A COMPANY	
5	CTGCTCCTCTGTTCGACAGTCAGCCGCATCTTCTTTTGCGTCGCCAG] (GTGAAGACGGGC	(INTRON A)	G:	TTCCATGGCACCGTCAAGGCTGAGAACGGGAAGCTTGTCATCAATGGAAATCCCATCACC	300
65	GGAGAGAAACCCGGGAGGCTAGGGACGGCCTGAAGGCGGCAGGGGGGGG	(ININON A)	f:	T G	
125	TGTGTTCGCGCCGCTGCGGGGTGGGCCCGGGCGGCCTCCGCATTGCAGGGGCGGGC		g:	PhoHisGlyThrValLysAlaGluAsnGlyLysLeuVallleAsnGlyAsnProlleThr	
185	GACGTGATOCGGCGCGGGCTGGGCATGGAGGCCTGGTGGGGGAGGGGA		•		
245	GTGTCGGCCGGGGCCACTAGGCGCTCACTGTTCTCCCTCC	IEXON 21	a ;	ATCTTCCAGGAGCGAGATCCCTCCAAAATCAAGTGGGGGGATGCTGGCGCTGAGTACGTC	360
305	GCTCAGACCATGGGGAAGGTGAAGGTCGGAGTCAACGG)	[5404 2]	g:	IlePheGlmGluArgAspProSerLysIleLysTrpGlyAspAlaGlyAlaGluTyrVal	
	(GTGAGTTCGCGGGTGGCTGG	(INTRON B)			
365	GGGGCCCTGGGCTGCGACCGCCCCGGAACCGCGTCTACGAGCCTTGCGGGCTCCGGGTCT		A :	GTGGAGTCCACTGGCGTCTTCACCACCATGGAGAAGGCTGGGGCTCATTTGCAGGGGGGA	420
425 485	TTGCAGTCGTATGGGGGCAGGGTAGCTGTTCCCCGCAAGGAGAGCTCAAGGTCAGCGCTC		£:	A C	
185 545	GGACCTGGCGGAGCCCGCACCCAGGCTGTGGCGCCCTGTGCAGCTCCGCCCTTGCGGCG CCATCTGCCCGGAGCCTCCTTCCCCTAGTCCCCAGAAACAGGAGGTGCCTACTCCCGCCC		g:	ValGluSerThrGlyVelPheThrThrMetGluLysAlaGlyAleHisLeuGlnGlyGly	
605	GAGATCCCGACCCGGACCCCTAGGTGGGGGACGCTTTCTTT		41	00011110000010101000000000000000000000	
665	TCACGTGTCGCAGAGGAGCCCCTCCCCCACGGCCTCCGGCACCGCAGGCCCCGGGATGCT		b:	GCCAAAAGGGTCATCATCTCTGCCCCCTCTGCTGATGCCCCCCATGTTCGTCATGGGTGTG	440
725	AGTOCGCAGCGGGTGCATCCCTGTCCGGATGCTGCGCCTGCGGTAGAGCGGCCGCCATGT		f:	A T	
785 845	TGCARCCGGGAAGGAAATGAATGGGCAGCCGTTAGGAAAGCCTGCCGGTGACTAACCCTG CGCTCCTGCCTCGATGGGTGGAGTCGCGTGTGGCGGGGAAGTCAGGTGGAGGAGGCTAG		a:	AlaLysArgVallleIleSerAlaProSerAlaAspAlaProMetPhevalMetGlyVal	
905	CTGGCCCGATTTCTCCTCCGGGTGATGCTTTTCCTAGATTATTCTCTGGTAAATCAAAGA		-		
965	AGTOGGTTTATOGAGGTCCTCTTGTGTCCCCTCCCCGCAGAGGTGTGGTGGCTGTGGCAT		A:	AACCATGAGAAGTATGACAACAGCCTCAAGATCATCAGCAATGCCTCCTGCACCAAC	540
1025	GGTGCCAAGCCGGGAGAAGCTGAGTCATGGGTAGTTGGAAAAGGACATTTCCACCGCAAA		f:	A TOGGT	
1005	ATGCCCCCTCTGGTGGTGGCCCCTTCCTGCAGCGGCTCACCTCACGGCCCGCCC		g:	AsnRisGluLysTyrAspAsnSerLeuLysIleIleSerAsnAlaSerCysThrThrAsn	
1145 1205	GGGTGTCTGGGGGCGCCTCGGGGAACCTGCCCTTCTCCCCATTCCGTCTTCCGGAAACCAGA		_		
1265	TCTCCACCGCACCCTGGTCTGAGGTCTGAGGTTAAATATAGCTGCTGACCTTTCTGTAGC		4:	TGCTTAGCACCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGAAGGACTCATG	600
1325	TGGGGGCCTGGGGCTCTCTCCCATCCCTTCTCCCCACACACA		g:	CysleuAlaProleuAlaLysValIleHisAspAsnPheGlyIleValGluGlyLeuMet	
1385	GCTCCCACTCCTGATTTCTGGAAAAGAGCTAGGAAGGACAGGCAACTTGGCAAATCAAAG			ACCACAGTCCATGCCATCACTGCCACCAGAAGACTGTGGATGGCCCCTCCGGGAAACTG	660
1445 1505	CCCTGGGACTAGGGGGTTAAAATACAGCTTCCCCTCTTCCCACCCGCCCCAGTCTCTGTC CCTTTTGTAGGAGGGACTTAGAGAAGGGGTGGGCTTGCCCTGTCCAGTTAATTTCTGACC		•:	C C	••0
1565	TITACICCICCCTITGAGTTTGATGATGCTGAGTGTACAAGCGTTTTCTCCCTAAAGGG		£:	T C	
1625	TGCAGCTGAGCTAGGCAGCAGCATCCTGGGGTGGCATAGTGGGGTGAATACC		q:	ThrThrValHisAleIleThrAlaThrGinLysThrValAspGlyFroSerGlyLysLeu	
1685	ATGTACAAAGCTTGTGCCCAGACTGTGGGTGGCAGTGCCCACATGGCCGCTTCTCCTGGA				
1745	AGGGCTTCGTATGACTGGGGGTGTTGGGCAGCCCTGGAGCCTTCAGTTGCAGCCATGCCT		A:	TGGCGTGATGGCCGCGGGGCTCTCCAGAACATCATCCCTGCCTCTACTGGCGCTGCCAAG	720
1805 1865	TARGECROGECAGECTGGCAGGGAGCTEAAGGGAGATAAAATTEAACCTCTTGGGCCCT CCTGGGGGTAAGGAGATGCTGCATTGGCCCTCTTAATGGGGAGGTGGCCTAGGGCTGCTC		f:	λ . д g	
1925	ACATATTCTGGAGGAGCCTCCCCCCCCCCCCCTCTTCTCTCCCCCCTCTTGTCTCTAG)		9:	TrpArgAspGlyArgGlyAlaleuGlnAsnIleIleProAlaSerThrGlyAlaAlaLys	
	[ATTTOG	[EXON 3]	4:		
1985	TCGTATTGGGCGCCTGGTCACCAGGGCTGCTTTTAACTCTGGTAAAGTGGATATTGTTGC		d:	GCTGTGGGCAAGGTCATCCCTGAGCTGAACGGGAAGCTCACTGGCATGGCCTTCCGTGTC AG	780
2045	CATCHATGACCCCTTCATTGACCTCAACTACATG {GTGAGTGCTACATGGTGACCCCCAAA	(INTRON C)	£1	Α	
2105	GCTGGTGTGGGAGGAGCCACCTGGCTGATGGGCAGCCCCTTCATACCCTCACGTATTCCC	(INTRON C)	q:	AlaValGlyLysValIleProGluLeuAsnGlyLysLeuThrGlyMetAlaPheArgVal	
2165	CCAG)		-		
	[GTTTACATGTTCCAATATGATTCCACCCATGGCAAATTCCATGGCACCGTCAAGGC	[EXON 4]	A :	CCCACTGCCAACGTGTCAGTGGTGGACCTGACCTGCCGTCTAGAAAAACCTGCCAAATAT	840
2225	TGAGAACGGGAAGCTTGTCATCAATGGAAATCCCATCACCATCTTCCAGGA) (GTGAGTGGA	(INTRON D)	t:	•••	
2285	AGACAGAATGGAAGAAATGTGCTTTGGGGAGGCAACTAGGATGGTGTGGCCTCCCTTGGGT	(THIRON D)	à:	ProThrAlaAsnVs1SerValValAspLeuThrCysArgLeuGluLysProAlsLysTyr	
2345	ATATGGTAACCTTGTGTCCCTCAATATGGTCCTGTCCCCATCTCCCCCCCACCCCGGTAG)			***************************************	
2405	[GCGAGATCCCTCCAAAATCAAGTGGGGCGATGCTGGCGCTGAGTACGTCGTGGAGTCCAC	(EXCN 5)	a: d:	GATGACATCAAGAAGGTGGTGAAGCAGGCGTCGGAGGGCCCCCTCAAGGGCATCCTGGGC	90D
2465	TGGCGTCTTCACCACCATGGAGAGGCTGGG] (GTGAGTGCAGGAGGGCCCGCGGGAGGGA	(INTRON E)	£:	C A	
2525	AGCTGACTCAGCCCTGCAAAGGCAGGACCCGGGTTCATAACTGTCTGCTTCTCTGCTGTA	(INIKON E)	g:	AspAspIleLysLysValVaiLysGlnAlaSerGluGlyFroLeuLysGlyFleteuGly	
2585	G)		•	, , , , , , , , , , , , , , , , , , , ,	
	[OCTCATTTGCAGGGGGAGCCAAAAGGGTCATCATCTCTGCGCCCCTCTGCTGATGCCCC	[EXON 6]	a:	TACACTGAGCACCAGGTGGTCTCCTCTGACTTCAACAGCGACACCCACTCCTCCACCTTT	960
2645	CATGITCGTCATGGGTGTGAAACCATGAGAAGTATGACAACAGCGTCAAGATCATCAG] (GTG	(INTRON F)	f:	с а тс	
2705	AGGAAGGCAGGCCCGTGGAGAAGCGGCCAGCCTGGCACCCTATGGACACGCTCCCCTGA	(ININON I)	g:	TyrThrGluHisGlnValValSerSerAspPheAsnSerAspThrHisSerSerThrPhe	
2765	CTTOCGCCCGGCTCCCTCTTTCTTTGCAG)				
	[CARTGCCTCCTGCACCAACTGCTTAGCA	[EXON 7]	à:	GACGCTGGGGCTGGCATTGCCCTCAACGACCACTTTGTCAAGCTCATTTCCTGGTATGAC	1020
2825	[CAATGGGTCGTGCAGGAGGAGGAGTGGTTAGGA		£:	AT T A T	1020
	[CANTGCCTCCTGCACGACGACTACTTTGCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGAAGGACTCATG] [GTATGAGAG	[EXON 7] {INTRON G}			1020
2885 2945	CCATOGCTCATGCACCACGACACTCTTACA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAAGGACTCATGI IGTATCAGA CTGGGGAATGGGATCGAGCTCCCACCTTTCTCTCTCACACACTGCTCTCCCTCTCGG CTGGGGAACGGCAACCTTGGGGGTTCTGGGGACTGGTTGTTCTTTCA		£:	A T T A T A A T A A T A AspAlaGlyAlaGlyIleAlaLeuAspAspAisPheVallySLeuIleSerTrpTyrAsp	
2885 2945 3005	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCACCACCACTCATCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGGACTCATGT GTATCAGGA CTGGGGAATGGGACTGAGGGTTCCCACCTTTCTCTCATCAAGACTGGCTCCTTCCT		f: g:	AT T A T	1020
2885 2945	CCATGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGGGGGACGACTCATTGCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGGAGGACTCATGI [GTATCAGG CTGGGGAATGGGATCGCAGCTTTGCTGTTCCATCAAGGACCTCGTGTGGG GCTGCGTGCAACCCTTGGGGTTGTGGGGTTCTGGGGATTGCTGTTCCATTAATTTCCTTTCA AGGTGGGGAGGGGGGGGTAGAGGGGTAATGGGGAGTAGCTGCGTGAGGGCCCACTCCTTTT GCA1]	(INTRON G)	f: g: a: f:	A T T A T A ASPALAGIYALAGIYILAALAUAANAASPALASPALAGIYALAGIYILAALAUAANAASPALASPALAYSLAUILASATTIFTTYTAASPAACAGGATGGTGGTGGACCTCATGGCCCACATGGCCTCCAAGGAG T A A	
2885 2945 3005	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCACCACCACTCATCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGGACTCATGT GTATCAGGA CTGGGGAATGGGACTGAGGGTTCCCACCTTTCTCTCATCAAGACTGGCTCCTTCCT		f: g:	A T T A T T A T T A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T T A T T T A T T T T A T T T T T T T T T T T T T T T T T T T T	10#0
2885 2945 3005 3065	CCATGGCCAAGGTCATCGATGACAACTTTGGTATCGTGCACCACCACTCATTCCACCACCACCACTCCTTACCACCAC	(INTRON G)	f: g: a: f: g:	A T T A T T A T T A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T T A T T T A T T T T A T T T T T T T T T T T T T T T T T T T T	
2885 2945 3005 3065 3125 3185 3245	CCATGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCAACTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGGAAGGACTCATGI GTATCAGAG CTGGGGAATGGGACTTGGGGTTCTGGTATCAAGGACTGCCTCCCCGTGTGGG GCTGCGTGCAACCCTTGGGGTTCTGGGGATTGCGTTCCATTAATTTCCTTTCA AGGTGGGGAGGGTAAGGGGGCTTGATGGGGATTCCGCTCCAGGGGCCTCCTTTG CCAG] [ACCACGTCCATCCATCACTCCCCCCCCCTGCAGGGCCTCCTCTTGCTGTGATGGCCCTCCCGGAACTGCGCGCTCCCTCC	(INTRON G)	f: g: a: f: g:	A T T A T T A T T A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T A T T T A T T T A T T T A T T T T A T T T T T T T T T T T T T T T T T T T T	10#0
2885 2945 3005 3065 3125 3185 3245 3305	CCCTGGCCAAGGTCATCCATGACAACTTGGTATCGTGCACCACCACTCTTTCCACCACCACCACTCGTTACCACCACCACTCGTTACCATGACAACTTGGTATCGTGCAAGGACTCATGTTGATCAGAGACTCAGTTGCTGCAGCACCACTCTCTCATCCAAGACTGCCTCCCTC	(INTRON G)	f: g: a: f: g: a: f:	A T AspalaciyalaciyileAlaLeuAshaspkisPheVallysLeuIleBerTrpTyrasp AACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCACATGGCCTCCAAGGAG T A AshGluPheGlyTyrBerAshArgValValAspLeuMetAlaHisMetAlaSerLysGlu TAAGACCCCTGGACCACCAGCCCCAGCAACAGCACAAGAGGAAGAGAGAGAGACACCTCACTG T	10#0
2885 2945 3005 3065 3125 3185 3245	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCACCACCAACTGCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGGACGACTCATGI CTGGGGAATGGGACTGAGGGTCCACCTTTCTCATCGAAGACTGGCCCCCCCC	(INTRON G)	f: g: a: f: g: a: f:	A T ApA1 at y at a transparent and transparent	10#0
2885 2945 3005 3065 3125 3185 3245 3305 3365 3425	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCGCCACCACCACTCCTTCACCACCACCACCACCACTCCATGCAACTTTGGTATCGTGCAGGCACCACCACGACGACCACGACGACCACGACGACCACGACG	(INTRON G)	f: g: a: f: g: a: f:	A T AspalaciyaladiyileAlaLeuAshaspkisPheValtysLeuIleBerTrpTyrasp AACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCAACAGGAG T AshaluPheGlyTyrBerAshArgValValAspLeuMetAlsHisMetAlsFerLysGlu TAAGACCCCTGGACCACCAGCCCCAGCAACAGCACAACAGGAAGAGAGAGACCCTCACTC T CTGGGGAGTCCCTGCCACACTCAACTCCACCCACCACACTGAATCTCCCCTCCTCACAGTTG A	10#0
2885 2945 3005 3065 3125 3185 3245 3305 3425 3485	CCCCTGGCCAAGGTCATCGATGACAACTTCTACCACCACCACCACCACCACTCCTACCACCA	(INTAON G)	f: g: a: f: g: a: f:	A T ApA1 at y at a transparent and transparent	10#0
2885 2945 3005 3065 3125 3185 3245 3305 3365 3425	CCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCGCCACCACCACTCTTCACCACCACCACTCCTTCCACCACCA	{INTROM G} [EXON 8] [INTROM H]	f: g: a: f: g: a: f:	A T AspalaciyalaciyileAlaLeuAshAspaisPheVallysLeuIleBerTrpTyrAsp AACGAATTTGGCTACAGCAACAGGGTGGTAGACCTCATGGCCCACACAGGAG T AssGluPheGlyTyrBerAshArgValValAspLeuMetAlaHisMetAlaSerLysGlu TAAGACCCCTGGACCACCAGCCCCAGCAACAACAACAAAAGGAAGAAGAA	10#0
2885 2945 3005 3065 3125 3185 3245 3305 3365 3425 3485 3545	CCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCACCACTCCTTCCATGCA CTGGGCAAGGTCATCCATGACAACTTTGGTATCGTGAAGGACTCCATGG CTGGGAATGGGACTCAGGGGTTCAGGGGACTCCATCCATTATTTCCTTTCA AGGTGGGGAGGGGGTGAGGGGGGGTAGGTGGGGGAGAGGGGCTCACTCCCTTTCCATAATTTCCTTTCA GCAGG (ACCACAGTCCATGCCATCCACCACCAGAGAGACTGTGAGGGCCTCACTCCGTGAA ACTGGGCGTGATGGGCGGGGGCTCTCCAGACAGACTCTTCCTTC	(INTAON G)	f: g: 	A T AspalaciyalaciyileAlaLeuAshaspkisPheValtysLeuIleBerTrpTyrasp AACGAATTTGGCTACAGCAACAGGGTGGTAGCCCCATGGCCCAATGGCCTCCAAGGAG T AspalaciyheGlyTyrBerAshArgValValAspLeuMetAlaHisMetAlaSerLysGlu TAAGACCCCTGGACCACCACGCCCCAGCAACAGCACAAGAGGAAGAGAGAGACCCTCACTG T CTGGGGGAGTCCCTGCCACACTCCACTC	1080 1140 1200
2885 2945 3005 3065 3125 3185 3245 3305 3425 3485	CCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCACCACTCCTTACCA CTGGGCAAGGTCATCCATGACAACTTTGGTATCGTGCAGGGACTCCATGGT CTGGGCAACCCTTGCTTCATCCAACACACCCTTCCCTT	{INTROM G} [EXON 8] [INTROM H]	f: g: 	A T A T A T A T A T A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A T A A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A	1080 1140 1200
2885 2945 3005 3065 3125 3185 3245 3305 3425 3485 3545 3665 3665 3725	CCCTGGCCAAGGTCATCCATGACAACTTGTCTACCACCACCACCACCACTCATTACCA CCCCTGGCCAAGGTCATCCATGACAACTTGTCATACTGCAAGGACCACTCATGTTGTCATGACAGGCTCATGTCTCATGACAACTTGTCATACATCATGCAAGGACCATGTTGTCATGACACCACTCATGTCATGACTCCCTCC	{INTROM G} [EXON 8] [INTROM H]	£; g: £; g: A: £; A: d: £; b: c: d:	A T AppliaGlyAlaGlyIiaAlaLauAnAsphisPheValtysLauIiaBerTrpTyrAsp AAGGAATTTGGCTACAGCACAGGGTGGTGGACCTCATGGCCCCACAGGAGG T AspGluPheGlyTyrBerAspArgValValAspLauMetAlaHisMetAlaSerLysGlu TAAGACCCCTGGACCACCAGCCCACAAAAAAAAAAAAAA	1080 1140 1200
2885 2945 3005 3065 3125 3185 3245 3305 3365 3425 3485 3545 3665 3765 3785	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCACCACTCCTTTCCATGCA CTGGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGGACACCACACGC CTGGCCAACCTTGCCTCCTTTCTCATCCAACACACCACCACCACACCACCACC	{INTROM G} [EXON 8] [INTROM H]	f: g: 	A T T ASPATAGLY AT T A T T ASPATAGLY	1080 1140 1200
2885 2945 3005 3005 3125 3185 3245 3305 3425 3485 3545 3665 3725 3785 3845	CCCTGGCCAAGGTCATCCATGACAACTTGTCTACCACCACCACCACCACTCATTACCA CCCCTGGCCAAGGTCATCCATGACAACTTGTCATACTGCAAGGACCACTCATGTTGTCATGACAGGCTCATGTCTCATGACAACTTGTCATACATCATGCAAGGACCATGTTGTCATGACACCACTCATGTCATGACTCCCTCC	{INTROM G} [EXON 8] [INTROM H]	£; g: £; g: A: £; A: d: £; b: c: d:	A T AppliaGlyAlaGlyIiaAlaLauAnAsphisPheValtysLauIiaBerTrpTyrAsp AAGGAATTTGGCTACAGCACAGGGTGGTGGACCTCATGGCCCCACAGGAGG T AspGluPheGlyTyrBerAspArgValValAspLauMetAlaHisMetAlaSerLysGlu TAAGACCCCTGGACCACCAGCCCACAAAAAAAAAAAAAA	1080 1140 1200
2885 2945 3005 3105 3185 3245 3305 3423 2485 3545 3665 3725 3785 3785 3845	CCCTGGCCAAGGTCATCCATGACAACTTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTCTTCTCATCCACCACCACCACCACTCTTTCCACCA	(INTRON G) [EXON 8] (INTRON N) (EXON 9)	f: g: 	A T T ASPAT AGINATED AS A T T ASPAT AGINATED AS A T AS	1080 1140 1200
2885 2945 3005 3125 3185 3245 3305 3425 3425 3465 3545 3665 3725 3865 3785 3865 3785 3865	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCACCACCACGCACCACCACGCACCACCACCACC	(INTRON G) [EXON 8] (INTRON H) (EXON 9)	f: g: 	A T Aspalaciyalaciyilealalauashasphisphevaltyslauilesertrptyrasp AACGAATTTGGCTACAGCAACAGGGTGGTAGCCTCATGGCCTCAATGGCCTCCAAGGAG T AspalaciyalaciytysserAsparqvalvalasplauMetalasismetalasertysgiu TAAGACCCCTGGACCACCACCCCACCAACAACAACAACAAAAGAAAAAGAAAAAGAACACCCCTCACCG T CTGGGGAGTCCCTGCCACACCTCAGTCCCCCACCACCACTGAATCTCCCCCTCCTCACAGTTG A T CCATGTAGAACCCCTTGAAGAGGGGAAGGGGAGCCTA-GGGAAGCGCACCTTGTCATGTA-CCATC T T C T C A C A	1080 1140 1260
2885 2945 1005 1065 1125 1185 3245 3305 3423 3423 3485 3545 3665 3725 3785 3785 3785 3785 3785	CCCTGGCCAAGGTCATCCATGACAACTTTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGCTATCCATGCACCACCACACTACTT CTGGGGAATGGGACTCATCCATGACAACTTTCTCTATCCAAGACTGACCTCCTTCCT	(INTRON G) [EXON 8] (INTRON H) (EXON 9)	f: g: f: g: a: d: f: d: f: d: f: d: f:	A T T ASPAT AGINATED AS A T T ASPAT AGINATED AS A T AS	1080 1140 1260
2885 2945 3005 3125 3185 3245 3305 3425 3425 3465 3545 3665 3725 3865 3785 3865 3785 3865	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCACCACCACGCACCACCACGCACCACCACCACC	(INTRON G) [EXON 8] (INTRON H) (EXON 9)	f: g: f: g: a: f: d: f: d: f: b: c: d: f: d: d: d: d:	A T Aspalaciyaladiyilealaleuashasphisshevaltysleuilesertrptyrasp AACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCCACTAGGCCTCCAAGGAG T Aspalaciyaladiyilealaleuashasphisshevaltysleuilesertrptyrasp AACGAATTTGGCTACAGCACAACAGGGTGGTGGACCTCATGGCCCCACCAACAGAGAGAAAAAGAGAAGAGAAGAGAACACACAGAGGAG	1080 1140 1260
2885 2945 3005 3065 3129 3185 3245 3305 3425 3425 3445 365 3725 3725 3785 3845 3965 3785 3845	CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACTCCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCACCACCACCACCACGCACCACCACCACCACCACCACCA	(INTRON G) [EXON 8] (INTRON H) (EXON 9)	f: g: a: f: g: a: d: f: d: f: a: d: c: d:	A T A T A T A T A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A T A T A T A T A T A T A T A T A T	1080 1140 1260
2885 2945 3005 3065 3129 3185 3245 3305 3425 3445 3545 3765 3772 3785 3785 3845 3965 4065 4065 4065 4075 4085 4085 4085 4085	CCCTGGCCAAGGTCATCCATGACAACTTGCTCATCGCACCACCACCACTCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTGCTATCTGCAAGGACCACTCATCTTACCA CTGGGGAATGGGACTGAGGGTCCCACCTTTCTCATCCAAGACTGGCTCCTCCTCCTCCGCGCGCCACCACCTTTCTCATCCAAGACTGCCCCCATCATTTCCTTTCA CTGGGCGAATGGACCACCGTGGGGGTTCTGCGCACCACCATTTTCCTTTCA AGCTGGCGAGGGGTGAAGGGGCTCACCCCACGAGAGACTGCGCCCCACGAGGCCTCACCCTCACCACCACCACCACCACCACCACCACCACC	(INTRON 6) (INTRON 1) (EXON 9) (3' FLANK)	f: g: f: g: a: f: d: f: d: f: b: c: d: f: d: d: d: d:	A T Aspalaciyaladiyilealaleuashasphisshevaltysleuilesertrptyrasp AACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCCACTAGGCCTCCAAGGAG T Aspalaciyaladiyilealaleuashasphisshevaltysleuilesertrptyrasp AACGAATTTGGCTACAGCACAACAGGGTGGTGGACCTCATGGCCCCACCAACAGAGAGAAAAAGAGAAGAGAAGAGAACACACAGAGGAG	1080 1140 1260
2885 2945 3005 3065 3129 3185 3245 3305 3425 3425 3445 3653 3725 3785 3845 3965 4025 4025 4085 4145 4095 4145 4145 4145 4145 4145 4145 4145 41	CCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCTTCACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGAAGGACCACACCACACTCCTTCACCAGCACCACCACGCACG	(INTRON 6) (INTRON 1) (EXON 9) (3' FLANK)	f: g: f: g: d: f: d: f: d: f: d:	A T A T A T A T A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T A T	1080 1140 1200 1260
2885 2945 3005 3065 3129 3185 3245 3305 3425 3425 3425 3425 3785 3655 4025 4085 4085 4145 4705 4705 4705 4705 4705 4705 4705 47	CCCTGGCCAAGGTCATCCATGACAACTTGTCTATCGCACCACCACCACTGCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTGTCATCATCGCACCACCACCACTGCTTGCCACCACCACGTCTCTCCATGATTCCTCATGACACCACGTGCACCACGTGCACCACGTTCTCCATGATTCCTTCC	(INTRON 6) [EKON 8] (INTRON H) (EKON 9) (3' FLANK)	f: g: f: g: d: f: d: f: d: f: d: f:	A T AspalaglyaladlylieAlaLeuashasphisPheVallysLeuIleBerTrpTyrAsp AAGGAATTTGGCTACAGCACCAGGGTGGTGGACCTCATGGCCCCACAGGAG T ASBGLUPheGlyTyrSerAsharqValVallaspLeuMetAlaHisMetAlaSerLysGlu TAMGACCCCTGGACCACCACCCCCACCACAACAGCACAACAGGAAGAGAGAGAGAGACCCTCACTG T CTGGGGAGTCCCTGCCACACTCACTCCCCCACCACCACCACCACCACCACCACC	1080 1140 1200 1260
2885 2945 3005 3125 3125 3165 3245 3305 3425 3425 345 365 3725 3785 3845 3965 4025 4025 4025 4705 FIG. 2.	CCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGCACCACCACCACTCCTTCACA CCCCTGGCCAAGGTCATCCATGACAACTTTGGTATCGTGCAGCACCACCACCACTCCTTCACAGGTCAACCTTTCCTTCATCCAAGACTTGCGTCACCACCACCACCACCACCACCACCACCACCACCACCAC	(INTRON 0) [EXON 0] (INTRON II) (EXON 9) (3° FLANK)	f: g: f: g: d: f: d: f: b: c: d: f:	A T AspAlaGlyAlaGlyIIeAlaLeuAshAspHisPhevallysLeuIIeSerTrpTyrAsp AACGAATTTGGCTACAGCAACAGGGTGGTGGTACCTCATGGCCCACATGGCCTCCAAGGAG T AspGluPheGlyTyrSerAshArgValValAspLeuMetAleHisMetAlaSerLysGlu TAAGAACCCCTGGACCACCACGCCCCACCAAACAGGAAGAAGAAGAAGAAG	1080 1140 1260 1260 1266
2885 2945 3005 3065 3129 3185 3305 3425 3425 3425 3425 3725 3725 3725 3725 3725 4085 4145 4725 4085 4145 4725 4085 4145 4725 4085 4145 4725 4085 4145 4725 4725 4725 4725 4725 4725 4725 47	CCCCTGGCCAAGGTCATCCATGACAACTTTGTCATCGACCACCACCACTGCTTACCA CCCCTGGCCAAGGTCATCCATGACAACTTTGTCATCATCACCACCACCACCACCACGTCATCATCACCACCACCACCACCACCACCACCACCACCAC	[INTRON 6] [EXON 8] [INTRON N] [EXON 9] (3* FLANK) Immediate hypmerase ional start	f: g: c: f: g: d: f: f: d: f: f: d: f:	A T ASPAIRAGLYALAGIYILANIALAUANAAPHISPHAVALLYSLAULISPOTTRTYTASP AAAGAATTTAGCTACAGCAACAGGGTGGTGGACCTCATGGCCCCACAGGAGA AAAGAATTTAGCTACAGCACAACAGGGTGGTAGACCTCATGGCCCCACAGGAGA AAAGAATTTAGCTACAGCACCACACAGGGTGGACCACCACACACA	1080 1140 1280 1260 1266
2885 2945 3005 3065 3129 3185 3245 3305 3425 3425 3465 3725 3785 3845 3965 4025 4085 4145 4705 4166 4705 4167 4167 4167 4167 4167 4167 4167 4167	CCCCTGGCCAAGGTCATCCATGACAACTTTCTCATCGACCACCACCACCACTCATTCACA CCCCTGGCCAAGGTCATCCATGACAACTTTGTCATCAAGACCACACTCATCTTCACCACGACCACTCATCTCACAGGCTCCCTCC	(INTRON 0) [EXON 0] [EXON 0] (INTRON N) (EXON 9) (3° FLANK) Immediate lymerase ional start ignal), []	f: g: . a: f: g: a: f: d: f: b: c: d: f: FIG. 3 seque were c and M	A T AspAlaGlyAlaGlyIieAlaLeuashAspHisPheVallysLeuIieSerTrpTyrAsp AACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCCACAGGGCG T AspGluPheGlyTySerAshArgValValAspLeuMetAleHisMetAlaSerLysGlu IAAGACCCCTGGACCACCACCCCACCCCACAACAGGAAGAAGAAGAAGAA	1080 1140 1200 1260 1266
2885 2945 3005 3065 3129 3185 3245 3305 3425 3425 3465 3725 3785 3845 3965 4025 4085 4145 4705 4166 4705 4167 4167 4167 4167 4167 4167 4167 4167	CCCTGGCCAAGGTCATCCATGACAACTTGTATCOTGCACCACCACCACTCATTCACA CCCCTGGCCAAGGTCATCCATGACAACTTGGTATCOTGCAAGGACCACTCATGI GTATGAGA CTGGGGAATGGGACTCACGGGGGGCCCACTCTTCTATCCAAGACTGCCTCCCTC	(INTRON 0) [EXON 0] [EXON 0] (INTRON N) (EXON 9) (3° FLANK) Immediate lymerase ional start ignal), []	f: g: a: f: g: d: f: d: d: d: f: f: f: f: f: d:	A T AspalaglyalaglyiieAlaleuashasphisPhevallysleuileBertrptyrasp AAGGAATTTGGCTACAGCACAGGGTGGTGGACCTCATGGCCCCACAGGAG AAAGGAATTTGGCTACAGCACAGGGTGGTGGACCTCATGGCCCCACAGGAGGAG AAAGGAATTTGGCTACAGCACCACAGGGTGGTGGACCTCATGGCCCCAAGAGGAGGAAGACACACAGGAAGAGAAGAGAAGA	1080 1140 1260 1260 1266
2885 2945 3005 3065 3129 3185 3245 3305 3425 3425 3465 3725 3785 3845 3965 4025 4085 4145 4705 4166 4705 4167 4167 4167 4167 4167 4167 4167 4167	CCCCTGGCCAAGGTCATCCATGACAACTTTCTCATCGACCACCACCACCACTCATTCACA CCCCTGGCCAAGGTCATCCATGACAACTTTGTCATCAAGACCACACTCATCTTCACCACGACCACTCATCTCACAGGCTCCCTCC	(INTRON 0) [EXON 0] [EXON 0] (INTRON N) (EXON 9) (3° FLANK) Immediate lymerase ional start ignal), []	f: g: a: f: g: a: f: g: a: f:	A T AspAlaGlyAlaGlyIieAlaLeuashAspHisPheVallysLeuIieSerTrpTyrAsp AACGAATTTGGCTACAGCAACAGGGTGGTGGACCTCATGGCCCCACAGGGCG T AspGluPheGlyTySerAshArgValValAspLeuMetAleHisMetAlaSerLysGlu IAAGACCCCTGGACCACCACCCCACCCCACAACAGGAAGAAGAAGAAGAA	1080 1140 1200 1260 1266

FIG. 3. Comparison of the human GAPDH gene exon sequence to other GAPDH sequences.

The human GAPDH gene exon sequence (a) and lis deduced amino acid sequence (g), were compared to the gublished cDNA sequences for human myscle GAPDH (b) of Hanauer and Mandel (12), human liver GAPDH (c) of Iso, et al. (10), human liver GAPDH (d) of Acari, et al. (36), the human eriphticorys GAPDH (r) of Xifion, et al. (10), and a human GAPDH pseudogene (f) of Hanauer and Mandel (12), Only base differences with the human exon sequence or so thown. A buillet in the sequence () indicates a missing base Dashed lines (-) indicate that a sequence was not reported in these areas. The underlined sequences include the translational start site, the translational stop site, and the polyadenylation signal. Note bases 21 to 53 of sequence d are entered as the reverse complement of the published sequence (36).

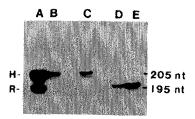


FIG. 4. Primer extension analysis of GAPDH mRNA.

Total RNA was isolated from human placental tissue, human JEG3 choriocarcinoma colls, rot 145 hopatoma colls transfected with pghGAPDH (731) and untransfected H35 hepatoma cells. GAPDH mRNA was quantilated by primer extension analysis as described in Experimental Procadules. The extension product for human GAPDH mRNA designated # (205 nucleotides) is 10 nucleotides longer than rat GAPDH mRNA designated # (195 nucleotides).

(205 nucleotides) is 10 nucleotides).

Lane A: 731 cells freated with 1 mu/ml insulin for 24 hrs prior to harvesting mRNA contained two predominant extension products, 205 and 195 nucleotides in length.

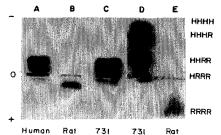
Lane B: Human placenta contained a single extension product of 205 nucleotides.

Lane C: Human JEGS choriocarcinoma cells contained a predominant extension product.

Lane C: Human JEG3 chonocarcinoma cells contained a predominant extension product of 195 nucleotides.

Lane D: H35 rat hepatoma cell contained a predominant extension product of 195 nucleotides.

Lane E: H35 rat hepatoma cells treated with 1 mu/ml insulin for 24 hrs prior to harvesting mRNA contained a predominant extension product of 195 nucleotides.



Human Rat 731 731 Rat

FIG 5. Isoenzyme analysis of GAPDH proteins.

A 14% starch gal was used to compare the histochemical activities and migration of GAPDH tetrameric enzymes. Cell extracts were matched for protein and were subjected to electrophorosis at pH 7.0 as described in Experimental Procedures. (c) indicates cathodo, (b) indicates origin, and (4) indicates anothe.

Lane B. 1435 hepatoma cell extracts designated Human.

Lane C: 1435 hepatoma cell extracts designated AH by DhGAPDH designated 731.

A 1% Seakem agarose gel was used to compare the histochemical activities and migration of GAPDH tetrameric enzymes (Lanes O and E) as described in Experimental Procedures. The agarose gel was aligned at its origin to that of the starch gels (Lanes AB, and C).

Lane D: 731 cell extracts display four distinct areas of GAPDH histochemical activity corresponding to the association of human (H) and rodent (B) suburits. Pure human GAPDH activity with four human subunits HHHH displayed predominant cathodic migration. Chimeric enzymes with increasing rodent subunit composition HHHR, HHAR, and HRRR displayed decreasing cathodic migration.

HUMAN GAPON GENE

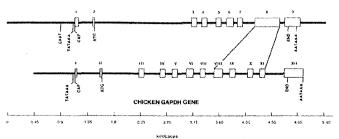


FIG. 6. Comparison of the human and chicken GAPDH genes.

The DNA sequence of the human GAPDH gene is shown above the chicken GAPDH gene. Open baxes indicate exons. The sequences are aligned at the putative transcription initiation site for each gene. Lines between the two genes indicate the loss of chicken introns to form a fused exon 8 in the human GAPDH gene.

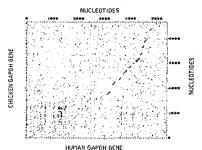


FIG 7. Dot Plot comparison of human and chicken GAPDH gene sequences. The chicken GAPDH gene (Y axis) and the human GAPDH gene sequence (X axis) were compared utilizing a Dot Plet program from the University of Wisconsin Genetics Group (23). Dots represent areas of 5 nucleotide sequence identity between the two genes.

TABLE |.

Comparison of sequences from the splice junction borders of the Human and the Chicken GAPDH Genes1.

		J EXON	a miren		a intron	2 Erou				
Exon No.		Pu#}	(GT	•	PyrAG)	(X 2-	*	Sequence	ident	ny3.
9.	49	DISTORDICACE	DEGAGACE		TOCOGCOCACE	EXCHANGE	44%)	165%	60%)	122%
£.	38	CLOCHOCHU	KIRAHOOOOO	•	TOCOCCAC	CANCACTAS	,,	(004	,	(200
2.	52	ACTICAACOOR	ETGACTTCG		GICICITAGE	PATTIDISTICO	100%)	189%	67%)	189%
H.	41	ACTICAADOO	proverege		TICTOCCAC	ATTTOOCCO	,	1	****	1417
3.	109	AACTACATOL	DEPORTAGE		TITOODOCACE	KITTIACATG	100%)	167%	44%1	[100%
191.	100	AACTACATO	DYDAGTIGE		CCCTCCTAC	CTTTACATG		10. 2	74.41	1,000
4.	107	CTTCCAGGA	KIROAOTOGA		00000000000	EXCENSION	100%)	(76%	93%1	[78%
IV.	104	CTTOCACION	GITGAGTAGCT	•	GITOCICAG	COMPACCOC	100-4)	1100	4471	1/474
5 .	9.1	ANGGCT0003	KITGAGIFOCA		CTOCTUTACE	KICTCATTTG	100%	156%	44%)	100%
₩.	#0	MOCCIOOO	KITAAGITAGG		TTTTTCAG	CICATORS		100 %	~~~;	185.4
6.	114	GATCATCACE	KOTOAGOAAS		TUTTTGCAGE	KAATGCCTC	£7%)	167%	56%)	[48%
VI.	116	AATTOTCAC	SHAMAKIEN	•	CTTTOCCAO	CANTOCATE	a	14. W	20 41	14.5.
7.	#2	CONCTCATOR	KRIATGAGAG		CTTTGCAG	BOCACAGIC	89%)	(58%	44%)	140%
Yn.	62	GOTOTTATO	KITACKITCTG		TITCCCTAC	ACCACITATE	****,	(***		(***
4.	413	CATTYCCTO	ESTATISTICAL		ATCTTCTACE	ESTATGAÇAA	78%1	(56%	33%1	(100%
VIII. (00)	413	tonricera	MATATOCACA	1	CONTRACT	PRINCE AND ADDRESS OF THE PARTY	1974)	100%	21.0	11000

<sup>The numan sequence is shown above the chicker sequence from Spore, et al. (40) which lacked a 3" flaving sequence, intervening nucleoides between the 5" and 3" ends of introns, (1" e sun borders, (1" e Indon borders
Consensus sequence at spice junctions for the numan and chicken GAPCH gence. Put = puther, type pyrimidine, X= any base of comparison, the chicken sequence has been displaced by one base region as displayed. For purposes of comparison, the chicken sequence has been displaced by one base to account for a possible making guaranter at the spice spite between exon M. Intron M., and exon M.</sup>